ISSN 2277-2685
IJESR/Oct-Dec. 2025/ Vol-15/Issue-4/62-73
Dnyaneshwari madhukar Sawant ***et. al.,*** **/ International Journal of Engineering & Science Research**

Crop recommendation system using Machine learning

[1]Dnyaneshwari madhukar Sawant, [2]Dr . Bijendra gupta
[1]Siddhant College Of Engineering, Sudumbre,
[2]Savitaribai Phule Pune University
Email - dnyaneshwarisawant9589@gmail.com

Abstract: This project focuses on developing an integrated system for agricultural optimization, including soil prediction, crop recommendation, plant disease detection, fertilizer suggestion, and crop yield prediction. The goal is to assist farmers in making informed decisions to improve agricultural productivity and sustainability through data-driven insights. The system leverages advanced machine learning and deep learning techniques to provide comprehensive support across various aspects of farming. The system begins by analyzing soil images to classify the soil type using a Convolutional Neural Network (CNN) model. Based on the identified soil type, it recommends suitable crops using Random Forest and XGBoost algorithms. The system also includes a plant disease detection module using a CNN model based on the MobileNet architecture. Farmers can upload leaf images, and the model identifies common diseases like blight, rust, and leaf spot. Early detection allows timely intervention, reducing crop loss and improving produce quality. For each disease, the system provides management strategies to prevent further spread. Fertilizer recommendations are made using Random Forest and XGBoost based on plant disease. Finally, the project employs LSTM to predict crop yield by considering various location, name of the area and soil parameters. This integrated approach assists farmers in making informed decisions for optimal crop selection, disease management, and maximizing agricultural productivity. Overall, the project demonstrates the potential of AI-driven solutions to address complex challenges in agriculture and contribute to global food security efforts. The project highlights AI's potential in tackling agricultural challenges, supporting sustainable farming practices, optimizing resources, and boosting productivity, for global food security.

Keywords – Soil analysis, Crop yield prediction, plant disease identification, CNN, Random Forest, XGBoost, and LSTM.

I. INTRODUCTION

The aim of this study is to develop an integrated system that leverages machine learning and deep learning techniques to optimize agricultural processes, including soil prediction, crop recommendation, plant disease detection, fertilizer suggestion, and crop yield prediction. The system is designed to assist farmers in making data-driven decisions that improve agricultural productivity, enhance sustainability, and reduce resource wastage

The **Next-Gen Groundwater Models** project aims to predict groundwater availability using advanced machine learning techniques. K-means clustering is first applied to the aqua dataset to categorize regions into clusters based on groundwater levels, identifying areas with varying water availability. Two classification models, Random Forest and Deep Convolutional Neural Networks (DCNN), are then utilized to predict water-bearing zones. Random Forest uses an ensemble approach for robust predictions, while DCNN captures intricate patterns in the data to enhance accuracy. The model's performance is evaluated using metrics like RMSE and MAE, and visualizations such as graphs and charts are generated to illustrate groundwater usage for both domestic and industrial purposes. The system is designed to support informed decision-making by testing and analyzing new data for real-time groundwater prediction.

The study focuses on integrating these machine learning and deep learning techniques into a unified system that can be used by farmers to optimize their farming practices. Additionally, the research emphasizes sustainability, ensuring that the system helps reduce resource wastage and minimize environmental damage.

Agriculture is a critical sector that sustains global food security, yet it faces numerous challenges including unpredictable climate conditions, soil degradation, plant diseases, and suboptimal resource management. To address these issues, advancements in artificial intelligence (AI) and machine learning (ML) offer innovative solutions for optimizing agricultural practices. This project introduces an integrated system aimed at revolutionizing traditional farming by providing data-driven insights for soil analysis, crop recommendation, plant disease detection, fertilizer suggestion, and crop yield prediction. By leveraging cutting-edge AI techniques, the system enables farmers to make informed decisions that enhance productivity and sustainability.

The core of the system begins with soil analysis, where a Convolutional Neural Network (CNN) model is utilized to classify soil types based on image inputs. Understanding the soil type allows for precise crop recommendations, further optimized using Random Forest and XGBoost algorithms. This data-driven approach ensures that crops are selected based on the most compatible soil conditions, boosting the likelihood of higher yields. Additionally, a plant disease detection module, also powered by a CNN model based on the MobileNet architecture, enables early identification of common diseases through image uploads of affected leaves. The system provides not only diagnoses but also actionable management strategies, allowing farmers to mitigate disease spread and reduce crop losses.

Beyond disease detection, the system offers tailored fertilizer recommendations that consider plant diseases, ensuring nutrient requirements are met efficiently. Crop yield prediction is another critical feature, implemented through Long Short-Term Memory (LSTM) networks that analyze location and soil parameters to forecast production levels. By integrating these various machine learning models, the system offers a comprehensive solution that addresses multiple facets of modern agriculture, contributing to more informed farming practices and improving overall food security.

II. LITERATURE SURVEY

Deep learning, a branch of artificial intelligence, has gained significant attention for its ability to automatically learn and extract features. It has been widely applied in various fields, including image, video, natural language processing and voice processing. In agriculture, particularly in plant protection, deep learning is increasingly used for plant disease recognition and pest assessment. Its application avoids the limitations of manual feature selection, making feature extraction more objective and improving research efficiency. This review highlights recent advancements in deep learning for crop leaf disease identification, discusses current trends, challenges, and aims to support researchers in plant disease and pest detection. [1]

This study reviews recent advancements in plant disease detection and classification using deep learning (DL) and machine learning (ML) models. By analyzing over 45 papers published between 2017 and 2020, it systematically evaluates state-of-the-art algorithms like SVM, Neural Networks, KNN, Naïve Bayes, AlexNet, and VGGNet. Each algorithm is compared based on processing techniques like image segmentation, feature extraction, and classification accuracy. The findings offer valuable insights for developing mobile-based applications that can help identify plant diseases and enhance agricultural productivity using data-driven approaches. [2]

Soil plays a crucial role in agriculture, with different types of soil possessing distinct features that affect crop growth. Understanding the characteristics of various soil types is essential for optimizing crop selection. Machine learning techniques offer valuable insights in this area and have seen significant progress in recent years. However, agricultural data analysis remains a challenging research field. In this paper, we propose a model that predicts soil series based on land type and recommends suitable crops accordingly. We utilize machine learning algorithms, including weighted Support Vector Machines (SVM), Bagged Trees, and k-Nearest Neighbor (k-NN), with a Gaussian kernel for soil classification. Experimental results demonstrate that the SVM-based method outperforms existing techniques, offering more accurate soil classification and crop recommendations. [3]

This research paper introduces a soil health intelligence system aimed at predicting soil type, pH levels, suitable crops, and nutrient content through image analysis techniques. By integrating image processing, deep learning, and machine learning, methods, the system analyzes soil images and uses RGB values to assess pH levels. The proposed model is trained with a soil image dataset using Convolutional Neural Networks (CNNs) and with pH-recognition data using regression models, with XGBoost Regressor yielding the best performance. The system demonstrates high accuracy and minimal error, showing great potential for real-world applications in agriculture. It can help farmers, agronomists, and researchers make informed decisions regarding soil management and crop productivity, contributing to sustainable agriculture practices. [4]

Agriculture is a vital part of the Indian economy, with over 50% of the population relying on it for their livelihood. However, environmental factors like weather and climate changes pose significant risks to agricultural sustainability. ML plays a key role in Crop Yield Prediction (CYP), helping farmers make decisions on what crops to grow and how to manage crops during the growing season. This research reviews various features and methods used in CYP, highlighting the limitations of techniques like Neural Networks and supervised learning, which often struggle with accuracy and nonlinear relationships. The study also explores ML models for crop yield estimation and classification, offering insights into improving agricultural productivity. [5]

This paper provides a comprehensive overview of deep learning applications in agriculture, highlighting various use cases such as crop classification, disease detection, and yield prediction. The authors discuss the potential of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) in analyzing agricultural data, emphasizing the importance of large datasets for effective model training. The survey identifies challenges such as data scarcity and the need for domain-specific models, offering insights into future research directions to enhance deep learning's impact on sustainable agriculture. [6]

Zhang and Wang review the role of machine learning in precision agriculture, focusing on its application in areas like soil health monitoring, pest detection, and crop management. The paper categorizes various machine learning techniques, including supervised and unsupervised learning, and discusses their effectiveness in real-time data analysis. The authors highlight case studies demonstrating improved agricultural practices through machine learning, while also addressing limitations such as the need for robust data integration and the challenge of model interpretability. [7]

This review by Whelan and Hakeem explores the integration of machine learning technologies in smart agriculture. The authors examine diverse applications, from yield prediction to automated irrigation systems. They emphasize the role of IoT and big data in facilitating machine learning solutions and discuss the implications for resource management and sustainability. The review also identifies barriers to adoption, such as the digital divide among farmers and the requirement for user-friendly interfaces to enhance accessibility. [8]

III. PROPOSED METHOD

In the proposed method, the methodology comprises three main components:

a) Soil Prediction and Crop Recommendation

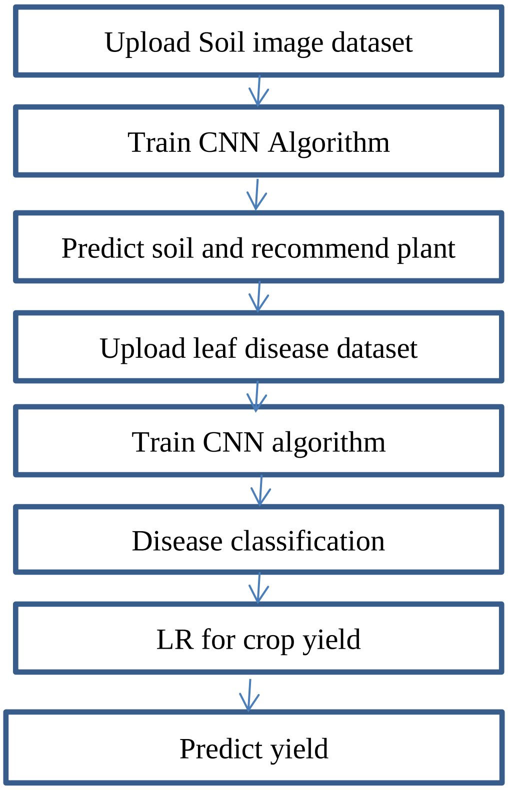
We start by analyzing soil images to classify soil types using a Convolutional Neural Network (CNN). This model leverages image data to provide accurate soil classification. Based on the identified soil type, we implement Random Forest and XGBoost algorithms to recommend suitable crops that can thrive in specific soil conditions. These algorithms utilize various features, such as soil composition and environmental factors, to optimize crop selection for improved yield.

b) Plant Leaf Disease Detection and Fertilizer Recommendation

A plant disease detection module is integrated, which utilizes a CNN model based on the MobileNet architecture. Farmers can upload images of plant leaves, and the model identifies common diseases, such as blight, rust, and leaf spot. Early detection enables timely intervention, reducing crop loss. For each detected disease, we recommend management strategies and appropriate fertilizers, leveraging the Random Forest and XGBoost algorithms to suggest the best options based on the disease and soil conditions.

c) Crop Yield Prediction

To predict crop yield, we utilize Long Short-Term Memory (LSTM) networks that analyze various parameters, including soil type, weather conditions, and crop history. This model is trained on historical data to understand patterns and forecast future yields accurately. By integrating these predictions with the earlier components, farmers receive comprehensive insights to make informed decisions about crop management and resource allocation.



System architecture:

4.1 Algorithms

❖ **Convolutional Neural Network (CNN):**

Used for soil prediction by analyzing soil images. The CNN model classifies soil types based on image data, helping in selecting the best crops for that particular soil type.

❖ **Random Forest and XGBoost:**

These ensemble learning methods are used for crop recommendation and fertilizer suggestion. After soil classification, Random Forest and XGBoost analyze soil characteristics and environmental factors to recommend suitable crops. They are also used to suggest fertilizers based on plant disease data.

❖ **MobileNet-based CNN:**

This is used in plant disease detection. Farmers upload images of plant leaves, and this CNN model identifies common plant diseases like blight, rust, and leaf spots, enabling early intervention to reduce crop loss.

❖ **Long Short-Term Memory (LSTM):**

LSTM networks are employed for crop yield prediction by analyzing historical data, soil parameters, and environmental factors. The model predicts future yields, assisting farmers in planning crop cycles and maximizing productivity.

❖ **K-Nearest Neighbor (KNN):**

Used for analyzing new data inputs for soil or crop classification based on historical data. KNN helps in identifying nearest matches for new crop data, aiding in decision-making.

❖ **Support Vector Machine (SVM):**

Used for classification of crops and plant diseases. SVM creates decision boundaries between different crop types or disease states, improving the accuracy of detection and recommendation systems.

4.2 Dataset:

For the dataset used in **Optimizing Agriculture Using Machine Learning**, the soil image dataset consists of **1281 images**, processed using feature extraction algorithms. The number of features before applying the extraction was **12,288**, and after feature extraction, **1200** key features were identified. The system leverages this processed data for tasks like soil prediction, crop recommendation, plant disease detection, and crop yield prediction.

Proposed Methodology for Agricultural Optimization

1. **Data Collection and Preprocessing**
   - **Soil Image Collection**: Gather soil images from various regions to ensure diverse data for training the CNN model.
   - **Data Cleaning**: Remove noise and irrelevant features from the dataset to improve model accuracy.
   - **Labeling**: Classify images according to soil types and corresponding metadata (e.g., geographic location).
2. **Soil Type Classification**
   - **Model Selection**: Use a Convolutional Neural Network (CNN) for image classification.
   - **Training the CNN**: Train the model on the preprocessed soil images, ensuring a split between training and validation datasets to avoid overfitting.
   - **Evaluation**: Assess model performance using metrics such as accuracy, precision, recall, and F1 score.
3. **Crop Recommendation System**
   - **Data Integration**: Combine soil type predictions with historical crop yield data and climatic conditions.
   - **Machine Learning Algorithms**: Implement Random Forest and XGBoost algorithms to recommend suitable crops based on the identified soil type and other environmental factors.
   - **Recommendation Output**: Generate a list of recommended crops for each soil type.
4. **Plant Disease Detection**
   - **Leaf Image Upload**: Allow farmers to upload images of plant leaves for analysis.
   - **CNN Implementation**: Utilize a CNN based on the MobileNet architecture for disease detection.
   - **Disease Classification**: Identify common diseases such as blight, rust, and leaf spot, and provide real-time feedback.
5. **Disease Management Strategies**
   - **Strategy Generation**: For each detected disease, provide farmers with management strategies, including preventive measures and treatment options.
   - **Integration with Crop Recommendation**: Link disease management strategies with recommended crops to ensure holistic management.
6. **Fertilizer Recommendation**
   - **Algorithm Application**: Use Random Forest and XGBoost algorithms to suggest appropriate fertilizers based on detected diseases and soil parameters.
   - **User Guidance**: Provide tailored fertilizer application rates based on specific crop and soil conditions.

7. **Crop Yield Prediction**
- **Data Collection**: Gather relevant parameters such as soil type, crop variety, and environmental conditions.
- **LSTM Model Development**: Train a Long Short-Term Memory (LSTM) model to predict crop yield based on historical data and current inputs.
- **Yield Forecasting**: Generate yield predictions that help farmers plan and make informed decisions regarding crop management.

8. **System Integration and User Interface**
- **Develop a User-Friendly Interface**: Create a mobile or web application that allows farmers to interact with the system easily.
- **Real-Time Feedback**: Enable features for real-time feedback and alerts based on the farmers' inputs and ongoing conditions.

9. **Testing and Validation**
- **Field Testing**: Validate the model outputs through field tests and gather feedback from farmers.
- **Performance Metrics**: Evaluate the system's effectiveness in improving productivity and resource optimization.

10. **Deployment and Continuous Improvement**
- **System Deployment**: Launch the application for broader farmer access.
- **Feedback Loop**: Establish a mechanism for continuous data collection and model improvement based on user feedback and changing agricultural conditions.

IV. RESULT

Results Analysis

In proposed method we used

a) Soil prediction and crop recommendation
b) Plant leaf disease detection and fertilizer recommendation
c) Crop yield prediction

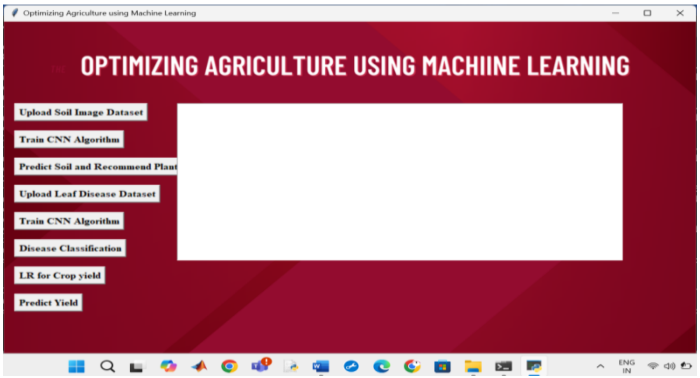


Fig. GUI for proposed Method

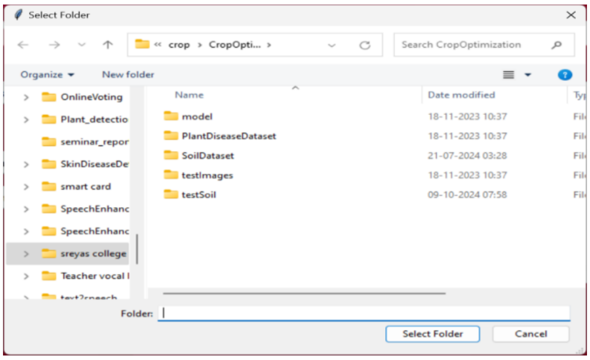


Fig. Upload the dataset related to soil

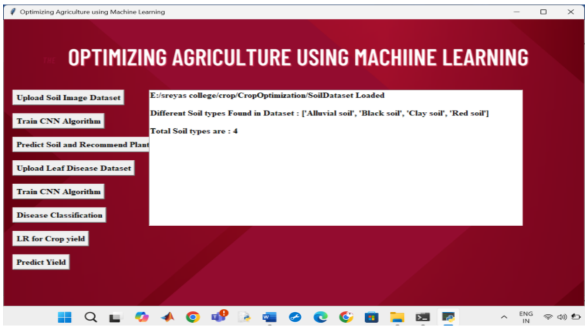


Fig. After uploading the soil dataset , details are shown

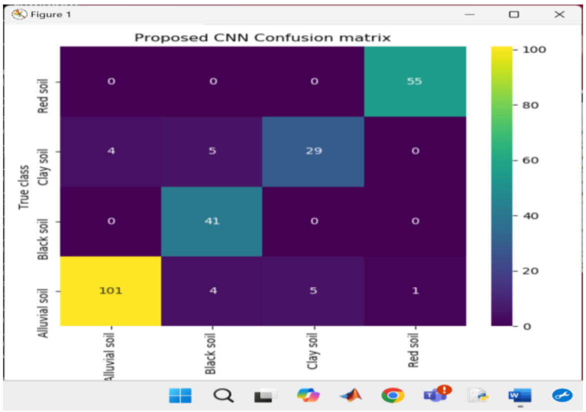


Fig. Confusion matrix for soil classification and plant recommendation

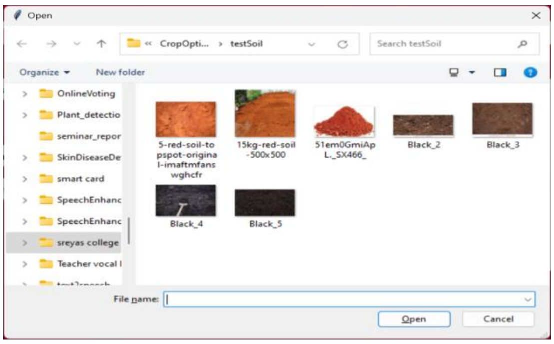


Fig. select soil image for testing

Fig. soil predicted and crop recommended

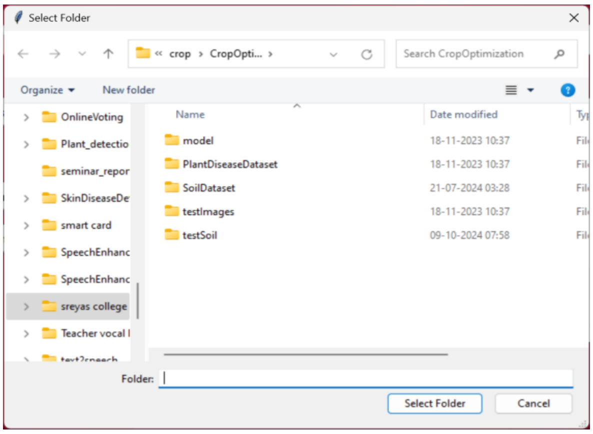


Fig. Uploading leaf disease dataset

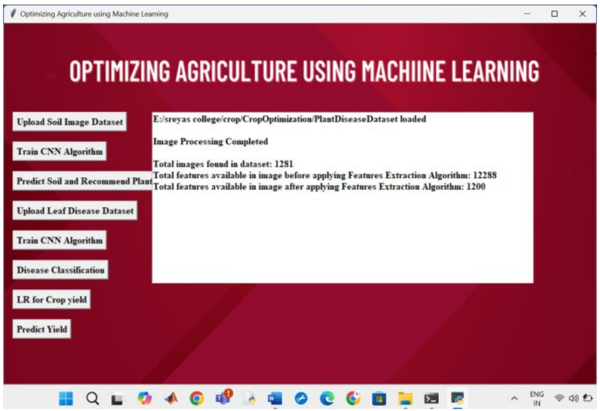


Fig. after uploading details are shown

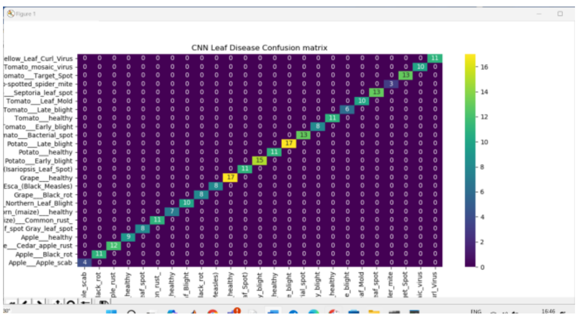


Fig. Confusion Matrix for plant disease detection and fertilizer recommendation

Normal leaf without disease

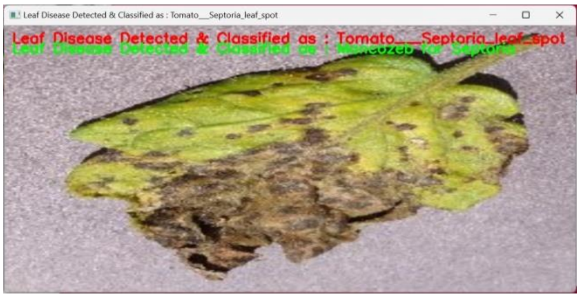


Fig. disease is predicted and fertilizer is recommended

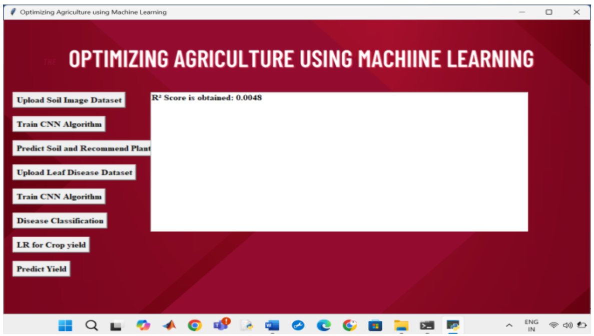


Fig. Linear regression R2 score is shown

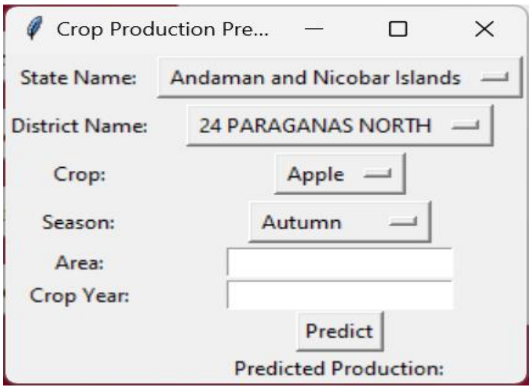


Fig. crop yield prediction interface

V. CONCLUSION

In conclusion, the paper demonstrates the transformative potential of AI-driven solutions in the agricultural sector. By leveraging machine learning models such as CNN for soil classification, Random Forest and XGBoost for crop recommendation and fertilizer suggestions, MobileNet-based CNN for plant disease detection, and LSTM for crop yield prediction, the system provides a comprehensive tool to assist farmers. This integrated approach supports sustainable farming practices, optimizes resource use, and enhances decision-making to improve crop productivity and reduce losses. Overall, this project highlights how machine learning can address key challenges in agriculture, contributing to increased productivity and global food security.

REFERENCES

1. Li, Lili, Shujuan Zhang, and Bin Wang. "Plant disease detection and classification by deep learning—a review." *IEEE Access* 9 (2021): 56683-56698.
2. Applalanaidu, Majji V., and G. Kumaravelan. "A review of machine learning approaches in plant leaf disease detection and classification." In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, pp. 716-724. IEEE, 2021.
3. Rahman, Sk Al Zaminur, Kaushik Chandra Mitra, and SM Mohidul Islam. "Soil classification using machine learning methods and crop suggestion based on soil series." In *2018 21st International Conference of Computer and Information Technology (ICCIT)*, pp. 1-4. IEEE, 2018.
4. Mali, Y. ., Rathod, V. U. ., Kulkarni, M. M. S. ., Mokal, P. ., Patil, S. ., Dhamdhere, V. ., & Birari, D. R. . (2023). A Comparative Analysis of Machine Learning Models for Soil Health Prediction and Crop Selection. *International Journal of Intelligent Systems and Applications in Engineering*, *11*(10s), 811–828. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/3335
5. D. J. Reddy and M. R. Kumar, "Crop Yield Prediction using Machine Learning Algorithm," *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2021, pp. 1466-1470, doi: 10.1109/ICICCS51141.2021.9432236.
6. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). "Deep Learning in Agriculture: A Survey." *Computers and Electronics in Agriculture*, 147, 70-90. DOI: 10.1016/j.compag.2018.02.016.
7. Zhang, C., & Wang, J. (2020). "Application of Machine Learning in Precision Agriculture: A Review." *Agricultural Systems*, 179, 102785. DOI: 10.1016/j.agsy.2019.102785.
8. Whelan, B., & Hakeem, A. (2019). "Machine Learning for Smart Agriculture: A Review." *Applied Sciences*, 9(19), 3937. DOI: 10.3390/app9193937.
9. Mohanty, S. P., Patil, A. S., & Barik, S. K. (2021). "A Comprehensive Review on Applications of Machine Learning in Agriculture." *Artificial Intelligence in Agriculture*, 4, 215-227. DOI: 10.1016/j.aiia.2021.08.001.
10. Liu, Y., Zhang, L., & Liu, Z. (2020). "Machine Learning in Agriculture: A Comprehensive Review." *Information Processing in Agriculture*, 7(1), 1-19. DOI: 10.1016/j.inpa.2019.09.002.
11. Xiong, Y., & Li, B. (2022). "Artificial Intelligence for Agriculture: A Review of Machine Learning Applications." *Computers and Electronics in Agriculture*, 194, 106757. DOI: 10.1016/j.compag.2022.106757.

12. Salehahmadi, Z., & Gharaviri, A. (2021). "Machine Learning for Smart Farming: A Review." *Agricultural Sciences*, 12(2), 218-233. DOI: 10.4236/as.2021.122017.

13. Mulla, D. J. (2013). "Twenty Five Years of Remote Sensing in Precision Agriculture: Key Advances and Remaining Challenges." *Remote Sensing*, 5(2), 618-646. DOI: 10.3390/rs5020618.

14. Schwalbenberg, A., & Crookston, R. (2020). "Machine Learning and Big Data in Agriculture: A Review." *International Journal of Agriculture and Biology*, 24(6), 1155-1160. DOI: 10.17957/IJAB/15.1201.

15. Zhang, Y., Wang, X., & Wang, Y. (2019). "Artificial Intelligence and Machine Learning in Precision Agriculture: A Review." *Journal of Integrative Agriculture*, 18(7), 1436-1450. DOI: 10.1016/S2095-3119(18)61941-1.